Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

Mon. Not. R. Astron. Soc. **402,** 1523–1535 (2010) doi:10.1111/j.1365-2966.2009.16017.x

The nature of H I absorbers in gamma-ray burst afterglows: clues from hydrodynamic simulations

Andrew Pontzen,[1⋆] Alis Deason,[1] Fabio Governato,[2] Max Pettini,[1] James Wadsley,[3] Thomas Quinn,[2] Alyson Brooks,[2,4] Jillian Bellovary[2] and Johan P. U. Fynbo[5]

[1]*Institute of Astronomy, Madingley Road, Cambridge CB3 0HA*
[2]*Astronomy Department, Box 351580, University of Washington, Seattle, WA 98195, USA*
[3]*Department of Physics and Astronomy, McMaster University, Hamilton, ON L8S 4M1, Canada*
[4]*California Institute of Technology, M/C 130-33, Pasadena, CA 91125, USA*
[5]*Dark Cosmology Centre, Niels Bohr Institute, University of Copenhagen, Juliane Maries Vej 30, DK-2100 Copenhagen, Denmark*

Accepted 2009 November 10. Received 2009 November 10; in original form 2009 August 25

ABSTRACT

In recent work, we have shown that it is possible to link quantitatively many aspects of damped Lyman α (DLA) absorbers in the spectra of quasars to high-resolution simulations of galaxy formation. Using runs from the same series of hydrodynamic numerical studies, we consider the expected properties of intrinsic Lyman α absorbers seen in the spectra of high-redshift ($z > 2$) gamma-ray burst afterglows (GRB–DLAs). If GRBs are associated with the death of massive stars, their afterglows provide insights into otherwise unprobed regions of protogalactic objects, but detailed physical interpretations are currently embryonic.

We find that median impact parameters (measured from the potential minimum) are approximately 1 kpc for GRBs compared with 4 kpc for quasi-stellar object–DLA (QSO–DLA). However, an equally important difference is that GRB–DLAs are predominantly associated with haloes of mass $10^{10} < M_{\rm vir}/{\rm M}_\odot < 10^{12}$, an order of magnitude larger than the hosts of QSO–DLAs. Accordingly, there are differences in the stellar properties of hosts. For instance, mean star formation rates are higher: $\langle \dot{M}_\star \rangle \simeq 10\,{\rm M}_\odot\,{\rm yr}^{-1}$ for GRB–DLAs compared with $\langle \dot{M}_\star \rangle \simeq 1\,{\rm M}_\odot\,{\rm yr}^{-1}$ for QSO–DLAs.

Our simulations accurately predict the form of the GRB–DLA H I column density distribution, producing quantitative agreement for $N_{\rm H\,I} > 10^{19}\,{\rm cm}^{-2}$, but they somewhat underpredict the incidence of low column densities $N_{\rm H\,I} < 10^{19}\,{\rm cm}^{-2}$. This is reflected in our estimate of the ionizing photon escape fraction, $f_{\rm esc} \simeq 1$ per cent, which is lower than the observational GRB-derived escape fraction (2 per cent). Line-of-sight neutral gas metallicities predicted by our simulations ($10^{-2} < Z/{\rm Z}_\odot < 1$) are consistent with the modest observational constraints. Because of large internal dispersions in gas metallicities, this agreement is not significantly compromised by imposing a cut-off on the metallicity of stars able to launch GRBs ($Z_\star < {\rm Z}_\odot/3$), confounding claims that the observed metallicity of GRB–DLAs poses a severe challenge to current GRB models.

Key words: galaxies: formation – quasars: absorption lines – gamma-rays: bursts.

1 INTRODUCTION

Observations of transient gamma-ray burst (GRB) afterglows have recently been established as a promising new window on the high-redshift Universe. In particular, the coupling of rapid response observing opportunities on large telescope facilities with the detection alerts provided by the *Swift* satellite (Gehrels et al. 2004) has produced a modest but growing sample of high-quality data. The GRBs are assumed (in the case of long-duration GRBs which have brighter optical afterglows and will be the focus of this work) to be associated with the death of Wolf–Rayet stars (MacFadyen & Woosley 1999).

Intrinsic absorption features in the GRB afterglow spectra are of particular interest because they probe the gaseous environment of the progenitor. The physical state of the gas may be quite

⋆E-mail: apontzen@ast.cam.ac.uk

different from that of typical (cross-section weighted) intervening absorbers. Dense neutral hydrogen giving rise to intervening absorption with column densities $N_{\rm H\,I}$ exceeding $2 \times 10^{20}\,{\rm cm}^{-2}$ is known as a damped Lyman α (DLA) system; the H I absorption and its associated metal line features are commonly studied in the spectra of quasars (e.g. Wolfe, Gawiser & Prochaska 2005). The focus of this paper will be the analogue of DLAs seen in the intrinsic spectra of GRB afterglows: neutral hydrogen absorption at the emission redshift of the GRB itself, by analogy termed as 'GRB–DLA'[1]. Because GRB–DLAs are physically associated with the GRB hosts, their nature may be quite different from intervening DLAs studied in quasar spectra (QSO–DLAs); part of the aim of this paper is to recap known observational differences and link them to the underlying protogalactic populations.

Indeed, a statistical comparison of QSO–DLAs with GRB–DLAs immediately confirms that the populations are distinct. While the frequency of quasar absorption systems decays monotonically with increasing neutral hydrogen column density (e.g. Wolfe et al. 2005; Prochaska et al. 2005), the observed frequency of GRB–DLAs grows to a peak at $N_{\rm H\,I} \simeq 10^{22}\,{\rm cm}^{-2}$ (a column density exceeding the strongest QSO–DLA ever observed; Jakobsson et al. 2006). Metallicities of the H I-associated gas, although challenging to measure in the transient GRB–DLA afterglows, are generally in the range $1/100 < Z/{\rm Z}_\odot < 1$, with a median somewhat exceeding ${\rm Z}_\odot/10$ (e.g. Fynbo et al. 2006; Savaglio 2006; Prochaska et al. 2007); whereas the median QSO–DLA metallicity is around ${\rm Z}_\odot/30$ (Pettini 2006).

The physical interpretation of the GRB–DLA population is maturing, but is not firmly established. It seems clear that, qualitatively, an association of GRB events with the death of massive, short-lived stars leads to the correct expectation of higher column densities (to ensure high star formation rates) and higher metallicities. This has led to the schematic picture (e.g. Prochaska et al. 2007) that GRB–DLAs preferentially probe the innermost regions of rapidly star-forming protogalactic haloes (Fig. 1), while QSO–DLAs intersect the outer regions of such systems. The neutral gas probed by the GRB–DLAs is not thought to be in the immediate circumstellar environment but rather at distances exceeding at least 10 pc. This follows because, under realistic assumptions, the ultraviolet (UV) emission from the afterglow itself is capable of ionizing any local neutral material within just an hour of the photons from the initial burst arriving (Prochaska, Chen & Bloom 2006). The non-local picture is reaffirmed by the non-detection of time variations in the neutral columns (Perna & Loeb 1998; Mirabal et al. 2002). In addition, approximate upper limits on distances of absorbing gas can be derived from UV pumping of fine-structure transitions, which suggest that the responsible neutral gas is at a distance of at least 100 pc (Prochaska et al. 2006) and perhaps as much as $\sim$1 kpc (Vreeswijk et al. 2007; D'Elia et al. 2009). Finally, dust depletion patterns seem more consistent with those seen in the warm neutral interstellar medium (ISM) of our Galaxy, rather than in cold dense molecular clouds (Savaglio 2006).

It is possible that, in addition to the GRB–DLAs probing particular regions of haloes, the population of objects giving rise to GRB–DLAs is on average more massive than those giving rise to QSO–DLAs. This will arise if the cold gas cross-section (relevant for QSO–DLAs) has a shallower than linear relationship with the UV luminosity (assumed relevant for GRB–DLAs); such a picture has been successfully employed by Fynbo et al. (2008) to reproduce differences in metallicities of the absorber populations.

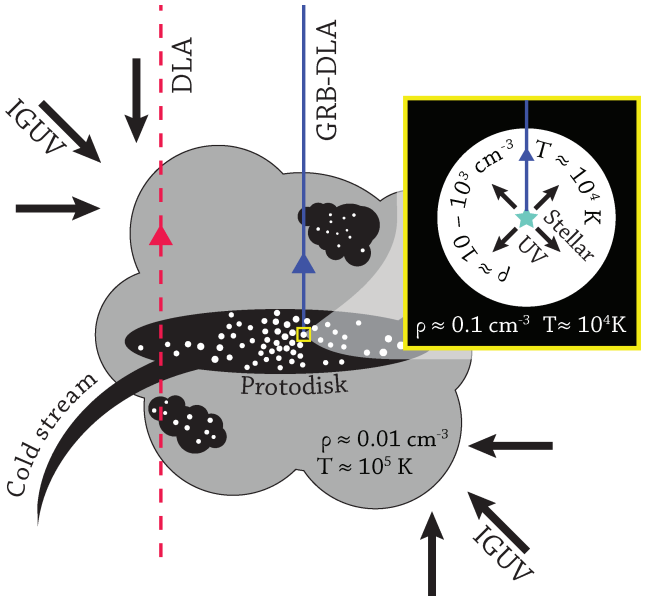

Figure 1. A sketch of relevant differences between typical QSO (dashed vertical line) and GRB afterglow (solid vertical line) sightlines through a protogalaxy at high redshifts ($z \sim 3$), after Prochaska et al. (2007). Assuming that GRBs are associated with the death of massive young stars, this class of object arises inside dense H II regions (white circles) within the most rapidly star-forming regions of the neutral medium (solid black regions); whereas QSO sightlines are more likely to intersect the neutral medium at a larger impact parameter. In our simulations, single protogalactic haloes (particularly those with $M_{\rm vir} \gtrsim 10^{10.5}\,{\rm M}_\odot$) can host multiple neutral star-forming regions; thus GRB sightlines may also intersect a separate neutral zone which has little physical association with their origin. The local stellar UV, presumed largely absorbed within the H II regions by our model (Section 2.1), has an effect on GRB–DLAs but appears unimportant in understanding the QSO–DLAs. However, the incoming background intergalactic mean field has an impact on both classes of object (Section 2.1). Typical temperatures and densities for the regions of interest are indicated (based on the observations of our own Galaxy).

As well as providing insights into individual objects, GRB–DLAs offer an unusually direct way to probe the escape of ionizing photons from the star-forming regions with which they are associated. On the assumption that the GRB rate traces the star formation accurately, one may construct a global escape fraction $\langle f_{\rm esc} \rangle \simeq 2$ per cent by modest extrapolation from the GRB–DLA column density distribution (Chen et al. 2007; Fynbo et al. 2009). Conversely, recent simulations by Gnedin, Kravtsov & Chen (2008), aimed at constraining escape fraction statistics, have incidentally been able to account for the GRB–DLA distribution (albeit with some mismatch at high column densities).

The emphasis in our work is reversed: we focus on statistics of GRB–DLAs between $2 < z < 4$, 90 per cent of which have column densities exceeding $10^{19}\,{\rm cm}^{-2}$. We use high-resolution hydrodynamical simulations with physically motivated star formation feedback algorithms (Stinson et al. 2006) to generate a catalogue of GRB–DLA sightlines. This catalogue can be compared to current observational constraints, while the association of each simulated sightline with a specific protogalactic object allows us to predict the relationship between GRB–DLAs and other known classes of objects (see also Fynbo et al. 2008). We can regard such associations with some confidence since, after calibrating two free parameters in isolated disc galaxy tests, the simulations are known to reproduce

[1]For simplicity, this term is applied to all intrinsic GRB absorption systems even though 20 per cent would not exceed the traditional DLA threshold.

realistic $z = 0$ disc galaxies (Governato et al. 2007, 2009), have realistic stellar mass–metallicity relations at low and high redshifts (Brooks et al. 2007; see also section 7.5 of Maiolino et al. 2008), and have produced a population of QSO–DLAs which are as realistic as current techniques permit (for a detailed comparison to observations see Pontzen et al. 2008, hereafter P08). The treatment of local ionizing sources we adopt is simple compared to detailed modelling employed in some ionizing photon escape fraction studies (e.g. Razoumov & Sommer-Larsen 2006; Gnedin et al. 2008; Wise & Cen 2009). As a result, the statistics of systems with $N_{\rm H\,I} < 10^{19}\,{\rm cm}^{-2}$ seem to be sensitive to assumptions in our approximations, and corresponding caution will be exercised in our interpretation – but results appear robust for larger column densities.

The simulations employed and techniques applied in this work are closely related to those of P08, and are discussed further in Section 2. By tracing sightlines from star-forming regions, we produce a set of simulated GRB–DLA statistics which we compare, where possible, with existing observations (Section 3). While both observational statistics and theoretical models are far less certain for GRB–DLAs than QSO–DLAs, we are able to draw tentative conclusions from our present study in Section 4.

2 SIMULATIONS AND TECHNIQUES

In this section, we describe the simulations employed and discuss the techniques required to build a cosmological sample of GRB–DLAs.

The simulations considered here are similar to those employed by P08. In particular, they take advantage of the 'volume renormalization' technique, in which the dynamic range is expanded by computing a subregion of the full volume at high resolution (Katz & White 1993): all results are taken from the high-resolution zone. Time evolution is computed with the smoothed particle hydrodynamics (SPH) code GASOLINE (Wadsley, Stadel & Quinn 2004) which implements gas cooling, the effects of a cosmological UV background (Haardt & Madau 1996) with a local approximate self-shielding algorithm (see section 5.4 of P08), and the effects of star formation based on the description of Stinson et al. (2006). Supernova ejecta abundances for oxygen and iron are calculated in the simulation following the yields of Thielemann, Nomoto & Yokoi (1986) and Weaver & Woosley (1993); after being injected into the interstellar gas these metals are advected with the SPH particles, although we now incorporate a correction for subresolution turbulent mixing which can transfer metals between Lagrangian fluid elements (see Shen, Wadsley & Stinson 2009). Gravitational softening is constant in physical units for $z < 8$ and evolves comovingly for $z > 8$; its final fixed value is given for each simulation in Table 1. For a full description, see section 2 of P08.

The particular regions in use for this work are listed in Table 1. These closely mirror the regions used by P08, although none is in fact identical. The first three are selected to form (by $z = 0$) individual objects and their immediate environments: 'DG' is a high-resolution run which, at $z = 0$, forms a dwarf galaxy ($M_{\rm vir} \simeq 3 \times 10^{10}\,{\rm M}_\odot$) at $z = 0$; 'MG' forms a Milky Way like galaxy ($M_{\rm vir} \simeq 7 \times 10^{11}\,{\rm M}_\odot$) and 'CL' forms a cluster ($M_{\rm vir} \simeq 1.1 \times 10^{13}\,{\rm M}_\odot$). 'C50' is a full cosmological box of side 50 Mpc (comoving). From each box we select, at $z = 3$, the set of haloes having greater than 3500 dark matter particles – an empirical limit which we established ensured convergence of our results including stellar populations (see also P08; Brooks et al. 2007). The 'CL' simulation, combining high resolution with large volume, contains 86 haloes satisfying this stringent criterion whereas the three other simulations have only ∼15 each. For this reason, we selected a random subsample of 15 haloes from the 'CL' simulation for our statistical results which are therefore derived from 60 haloes of high resolution, evenly distributed with $10^8 < M_{\rm vir}/{\rm M}_\odot < 10^{12.5}$. We discuss the impacts of the finite mass range of resolved haloes in Section 3.1.

As always with 'zoomed' simulations, one must be as careful as possible to test whether the arbitrary choice of region may have affected results through environmental dependences; we take agreement of statistics between boxes (where their resolved halo mass ranges overlap) as evidence that no severe problems are to be expected. In particular, our present simulations include at $z = 3$ an overdense protocluster region ('CL') which resolves the same range of halo masses as the more typical 'MG' box. There is no evidence for meaningful disagreement between the properties of haloes within these boxes (see also Crain et al. 2009). However, as in P08, we must accept, in principle, the possibility that we miss an environmental effect which has an impact on statistics – this is to be regarded as a trade-off against resolving such a large range of halo masses with relatively uniform resolution.

While the regions are from a different realization of the initial power spectrum than those presented in P08, the physics remains almost identical except for two notable differences. First, we have used the self-shielding approximation, described in section 5.4 of P08, in all our simulations. This has the effect of slightly raising cold gas fractions and lowering dwarf galaxy metallicities (the latter because the lower pressure of the ambient ISM makes our feedback algorithm somewhat more effective); see P08 for details. Secondly, we have implemented a turbulent metal diffusion algorithm as described at the start of this section.

Table 1. The simulations used in this work.

Tag	max $M_{\rm p,gas}({\rm M}_\odot)$	$M_{\rm p,DM}({\rm M}_\odot)$	ϵ (kpc)	Usable volume (comoving Mpc^3)
DG	$10^{3.9}$	$10^{4.6}$	0.12	4
MG	$10^{5.6}$	$10^{6.0}$	0.34	70
CL	$10^{5.6}$	$10^{6.0}$	0.34	600
C50	$10^{7.9}$	$10^{8.4}$	2.00	125 000

Note. The first column is the tag which we use to refer to each simulation. For all except 'C50', a subsample of the full box is simulated in high resolution; no results are taken from outside this region. The second and third columns refer, respectively, to the maximum gas and dark matter particle masses within the region, the fourth to the gravitational softening length (in physical units) and the final column gives the comoving volume of the region. The separate boxes are generated from entirely different sets of initial conditions; the 'DG' and 'MG' simulations are designed to form, respectively, a dwarf and approximately $L_\star$ galaxy at $z = 0$ while the 'CL' region forms a large cluster of galaxies at $z = 0$.

Because of the long lead-time involved in running the simulations considered, the cosmological parameters employed ($\Omega_{\rm M} = 0.24$, $\Omega_\Lambda = 0.76, \Omega_b = 0.042, h = H_0/(100\,{\rm km\,s^{-1}}) = 0.73, \sigma_8 = 0.77$, $n_s = 0.95$) are close to, but not fully, consistent with the latest constraints from large-scale observational results (e.g. Dunkley et al. 2009). As we have previously argued, we do not believe small differences in the adopted cosmology will impact strongly on the results (see section 5.3 of P08).

2.1 Approximate radiative transfer

Understanding the effect of radiation on the ionization state of the gas is crucial for obtaining realistic absorption statistics. The simulations incorporate a diffuse cosmological UV background based on Haardt & Madau (1996), uniform except in dense regions for which a local self-shielding algorithm, described in Section 5.4 of P08, is applied. We post-process the output with a simple equilibrium radiative transfer algorithm, described in Section 3.1 of P08, to improve our estimate of the self-shielding effect. We previously argued that the correction to this picture from local sources, which were not included, would be small.

However, in GRB–DLAs, where the sightline necessarily passes through highly ionized bubbles around hot stars, local radiation sources are likely to have a stronger impact on results (these hot bubbles presumably having negligible cross-section to a QSO sightline). Observationally, fine-structure transitions in GRB–DLA absorbers indeed suggest the presence of a local UV field (Prochaska et al. 2007).

A study of the local ISM shows that incorporating the stellar ionizing sources into a radiative transfer scheme using the raw density field from the simulations will overestimate the effect of local sources. This is because young stars are initially surrounded by a dense molecular cloud, which is rapidly photoevaporated to form an H II region. Such regions have high densities, typically between 10 and $100\,{\rm cm^{-3}}$ but sometimes as high as $10^4\,{\rm cm^{-3}}$. Hence, their recombination time is short and they are able to absorb the vast majority of the UV radiation produced within them. This microphysics occur on parsec scales, below the resolution limit of our simulations.

Assuming typical temperatures and densities for H II regions ($T = 10^4\,{\rm K}, n_{\rm H} = 30\,{\rm cm^{-3}}$), one may estimate the hydrogen mass required to fully absorb all UV photons produced, using an on-the-spot approximation to account for photons emitted by recombination to the ground state. Using STARBURST99 (Vázquez & Leitherer 2005) emission models, we find

$$M_{\rm H\,I\rightarrow H\,II} \simeq 1.5\times10^7\,{\rm M_\odot}\left(\frac{\dot{M}_\star}{1\,{\rm M_\odot yr^{-1}}}\right)\left(\frac{30\,{\rm cm^{-3}}}{n_{\rm H}}\right) \quad (1)$$

of neutral hydrogen would be ionized by the local sources, where $\dot{M}_\star$ is the star formation rate, the final factor is fixed by assumption in our calculations and a Kroupa, Tout & Gilmore (1993) initial mass function is assumed for consistency with the simulations.

To assess the impact of removing the corresponding neutral material, we calculate the total star formation rate per halo and expand ionized spheres around all young stellar particles ($<10^7$, yr) until the total mass converted from H I to H II is given by (1). Because the star-forming threshold density ($0.1\,{\rm cm^{-3}}$) in the simulations are lower than true H II region densities, the volume affected by this algorithm is rather larger than the physical sizes of H II regions would suggest, but by construction the mass affected is accurate. This should therefore reduce line-of-sight column densities by an appropriate fraction.[2] We verified that the impact on the quasar DLA statistics is, as expected, insignificant, but the effect on GRB–DLAs is non-negligible (Section 3).

Our approach ignores the effect of dust which is capable of absorbing a significant fraction of ionizing radiation (Osterbrock & Ferland 2006); since the quantity of dust in high-redshift H II regions is hard to assess, we merely admit that the magnitude of our correction could be reduced by including dust (see also Section 4). We further fail to account for molecular clouds, which could be important in understanding DLAs and GRB–DLAs (e.g. Schaye 2001). The small fraction of detectable H_2 in most QSO–DLAs (Ledoux, Petitjean & Srianand 2003) and lack of H_2 in many GRB–DLAs (Tumlinson et al. 2007; although see Fynbo et al. 2006; Prochaska et al. 2009) do not imply that H_2 is unimportant, as large masses could conceivably reside in small clouds with negligible cross-section. Ledoux et al. (2009) also suggested that non-detections in most GRB spectra are expected given the low metallicities of typical systems observed with high-resolution spectroscopy. Overall the fraction of gas in H_2 clouds is, at present, not calculable from our simulations and so we are forced to proceed on the assumption that later inclusion will make only a small correction to our results. Further discussion of the inherent problems in modelling the small-scale ISM and its impact on our results is given in Section 4.

2.2 Sightline generation

The theoretical and observational picture of the precise origins of long GRBs is complex (e.g. Woosley & Bloom 2006 and references therein). We will suppose the events to be associated with the death of massive stars ($\sim$25 $\rm M_\odot$); such stars have a main-sequence lifetime of just a few million years, far smaller than the dynamical time of host haloes (e.g. for the adopted cosmology at $z = 3$, $\tau_{\rm dyn} = (200\,G\rho_{\rm c})^{-1/2} \simeq 7\times10^8$ yr). For the purposes of this work, long GRBs are therefore assumed to arise in a given region with probability proportional to the star formation rate of that region. We choose the star particles younger than 5×10^7 yr in each halo as launch sites for our tracer sightlines in order to reflect this assumed association. Clearly the maximum age exceeds by an order of magnitude the actual lifetime of the GRB progenitors, but it is required to guarantee at least 200 candidates launch sites in each star-forming halo (in turn required for reliability of our resampling mechanism described in Section 2.3). We verified that reducing cut-off times to 10^7 years (the minimum possible at our resolutions) did not significantly affect our results with the exception of the low column density distribution ($<10^{19}\,{\rm cm^{-2}}$); see Section 3.2.

Additionally, we considered the effect of a metallicity ceiling for progenitor stars; such a limit is motivated by GRB collapsar models which require mass loss over the progenitor Wolf–Rayet stellar lifetime to be small (MacFadyen & Woosley 1999). However, we found its effect – even on the measured column metallicities – to be minor (a discussion is given in Section 3.6).

The sightline properties are generated using the code described in P08, performing a full SPH line-of-sight integration to determine the column densities of neutral hydrogen and neutral-phase metals,

[2] There is some evidence to suggest that the GRBs themselves ionize material along their line of sight (see the Introduction); with our current knowledge this is hard to model in detail, but we discuss possible consequences for our results in Section 4.4.

i.e. the sightline metallicity is defined to be

$$Z_{\rm sightline} = \frac{\int_0^\infty {\rm d}x\, Z(x)\, n_{\rm H\,I}(x)}{\int_0^\infty n_{\rm H\,I}(x)}, \quad (2)$$

where x is the distance from the GRB along the line of sight and $Z(x)$ is the metal mass fraction at that point. Equation (2) makes the implicit assumption that observers measure metal ions which are perfectly coupled[3] to the H I (e.g. for silicon, $n_{\rm Si\,II}/n_{\rm Si} = n_{\rm H\,I}/n_{\rm H}$).

We use the oxygen mass abundance reported by the simulations as our metallicity tracer, using the reference value $12 + \log_{10}(n_{\rm O}/n_{\rm H})_\odot = 8.69$ (from Lodders 2003) to normalize to a solar metallicity scale. The choice of oxygen reflects its status as the most abundant metal by mass; differences between elements are small relative to the overall spread of metallicities under consideration.

2.3 Resampling

The set of haloes extracted from all boxes is used to generate a cosmological sample in a directly analogous fashion to that described by P08. However, instead of using a gaseous cross-section we arrange that the probability of selecting a sightline from a particular halo of virial mass $M_{\rm h}$ is proportional to the product of the halo mass function $f(M_{\rm h})$ and the star formation rate of that halo $\dot{M}_{\star,{\rm h}}$. We bin our ensemble of haloes in virial mass; the probability of a given GRB originating in the halo h from mass bin $i(h)$ is then proportional to $w_{\rm h}$,

$$w_{\rm h} = \frac{F_{i(h)}\dot{M}_{\star,{\rm h}}}{n_{i(h)}}. \quad (3)$$

where F_i is the halo mass function integrated over the mass range of bin i (P08, equation 4) and n_i is the total number of haloes from our simulations in that bin. All equations from P08 follow identically with our replacement definition (3) for the halo weight $w_{\rm h}$.

3 RESULTS

3.1 Origins of GRB–DLAs

Because GRB–DLAs are intrinsic to the galaxy hosting the GRB event itself, the cosmological regions probed by such absorbers are not necessarily the same as those probed by the population of intervening QSO–DLAs. By assumption, our method triggers GRBs in a given region proportionally to that region's contribution to the cosmological star formation rate. This can be split into two factors: the relative contribution of haloes of different masses and the position of the star-forming regions within each halo. In this section, we will consider each of these two contributing effects in turn.

Fig. 2 shows the product of the halo mass and star formation rate $[f(M_{\rm h})\,\dot{M}_{\star,{\rm h}}]$ for individual haloes in each of our boxes; as described in Section 2.3, this relationship defines the relative contribution of haloes of the specified virial mass. We find that a broad range of haloes contribute to the overall population of GRB hosts,

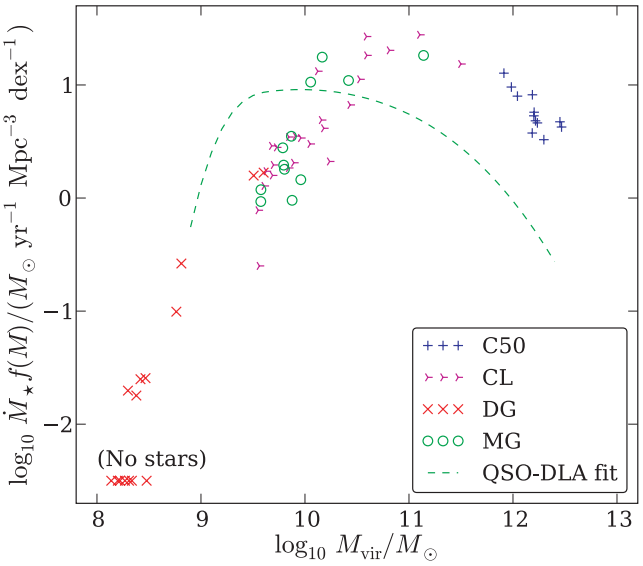

Figure 2. The product of the star formation rate and halo mass function for each qualifying halo (see Section 2 for resolution criteria) in our DG (crosses), MG (circles), CL (tripods) and C50 (plus symbols) runs. This quantity is proportional to the contribution of a given halo mass to the overall cosmological ensemble of GRBs and therefore GRB–DLAs. A fit to the equivalent quantity for quasar DLAs is shown as a dashed line, based on the locus of points in fig. 5 of P08. Haloes with no star formation are grouped together at the bottom of the figure.

peaking at $\sim 10^{11}$ M$_\odot$. The distribution is in good agreement between our four separate boxes in their regions of overlap. In particular, the proto-cluster (CL) environment shows no systematic differences from the disc galaxy progenitor region (MG).

For masses $M_{\rm vir} < 10^{9.5}$ M$_\odot$, the contribution tails off as star formation rates are stifled by the stellar feedback and UV field (e.g. Read, Pontzen & Viel 2006, P08). The resulting steep drop at low masses in Fig. 2 shows that our low-mass cut-off of $M_{\rm vir} = 10^8$ M$_\odot$ should not lead to any systematic effects. For $M_{\rm vir} > 10^{11.5}$ M$_\odot$ the GRB rate contribution begins to be suppressed by the downturn in the cosmological halo mass function; however, somewhat unsatisfactorily, our most massive halo has $M_{\rm vir} = 10^{12.5}$ M$_\odot$ which means that the volume of our simulations is insufficient to trace objects until they are so rare as to be cosmologically insignificant. We checked for systematic effects by artificially cutting off our distribution at 10^{12} M$_\odot$, finding that no results were severely affected, but a satisfactory resolution would require considerably larger simulations to produce a full population of rare $M_{\rm vir} > 10^{12.5}$ M$_\odot$ objects.

We have overplotted (dotted line in Fig. 2) a fit to the equivalent halo-weighting quantity for QSO–DLAs, $\sigma_{\rm DLA}\ f(M)$, where $\sigma_{\rm DLA}$ is the cross-section of the halo for column densities of neutral hydrogen $N_{\rm H\,I} > 10^{20.3}$ cm^{-2} from P08. The dominant contribution to the QSO–DLAs can be seen to arise in haloes approximately 10 times less massive than for GRB–DLAs; the mechanisms determining the upper and lower limits are similar in both cases, but because $\sigma_{\rm DLA}$ scales less steeply than $\dot{M}_\star$ ($\sigma_{\rm DLA} \sim M_{\rm vir}$, P08 fig. 4; $\dot{M}_\star \sim M_{\rm vir}^{1.6}$, P08 fig. 13), the latter distribution is skewed towards higher masses. We will comment in Section 3.5 that this has important consequences for the metallicity distribution of our simulated GRB–DLAs.

We now turn to the location of GRB–DLAs within individual haloes. Fig. 3 shows the impact parameter b (projected distance

[3]This assumption was also made for generating Si II profiles in P08; however, contrary to the statement in Tescari et al. 2009, we did not assume a fixed metal *abundance* when generating Si II profiles. We also used CLOUDY models to show that the correction to the perfect coupling model is small (P08, section 5.2).

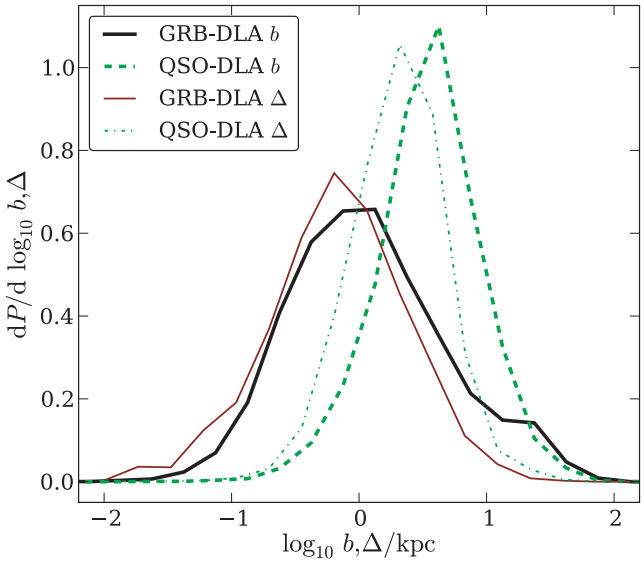

Figure 3. The distribution of impact parameters b for GRB–DLAs (solid thick line) and QSO–DLAs (dotted thick line). The GRB–DLA population typically originates very close to the centre of the dark matter halo, whereas the QSO–DLAs tend to pass several kpc from the centre, although in both cases there is a long tail to large impact parameters. This tail is associated with massive ($M_{\rm vir} > 10^{11}\,\mathrm{M}_\odot$) haloes with recent merging and hence multiple gas/star formation centres. The longitudinal distances, Δ, over which the absorbers are spread (see the text for exact definition), are also shown as thinner lines; the GRB–DLAs are seen to be more compact along the line of sight. All units are physical.

of the DLA sightline to the minimum gravitational potential of the host halo) for GRB–DLAs (thick solid line) and QSO–DLAs (thick dashed line) in our current set of simulations. The former arise considerably closer to the halo centre (typically at ∼1 kpc) than the latter (at ∼4 kpc; although there is a spread over orders of magnitude in distance for both cases). The most active star formation is naturally associated with the deepest part of the potential well, so that these results can be interpreted as confirming the intuitive picture of Fig. 1.

In both the QSO and GRB case there is a tail of impact parameters extending to 100 kpc distances, arising in high-mass haloes ($M_{\rm vir} > 10^{11}\,\mathrm{M}_\odot$) which are dynamically young and contain multiple H I clouds hosting active star-forming regions. Thus, care is needed in interpreting the highest b-values; even though in each case the brightest star-forming regions are associated with the potential minimum, observers could expect to observe significant star formation closer to the GRB sky coordinates than these extreme values would at first suggest. In addition, small haloes have such slow star formation that they may be invisible for practical purposes.

Furthermore, some caution should be exercised in interpreting b-values smaller than a kiloparsec, since these lie below the gravitational softening scale of the coarsest simulation included (Table 1). Reassuringly, however, we found that the only effect of removing the low-resolution C50 box from our sample was to truncate the b-value distribution above approximately 20 kpc, simply reflecting the removal of all high-mass haloes from our sample: no change in the low-b tail was produced.

To compare the gas distribution along the sightline of intrinsic and intervening absorbers, we define Δ to be the distance over which the central half of the H I column is assembled (i.e. the physical distance between lower and upper quartile positions). The distribution of this quantity is overplotted in Fig. 3 as thin lines (solid for GRBs, dashed for QSOs). Δ values suffer equally from the resolution caveat mentioned for b above; however, we can nonetheless conclude that GRB–DLAs are typically assembled over a longitudinal distance somewhat less than a kiloparsec whereas QSO–DLAs are built up over larger distances (∼2 kpc). This indicates that the H I columns in our typical GRB–DLAs are not assembled from the immediate environment of the GRB itself, but rather from gas spread through kiloparsec-sized surrounding regions.

Although the distributions of b and Δ are seen to be similar to each other, we found that for individual absorbers these two quantities are in fact almost uncorrelated, i.e. there is no significant relation between the longitudinal dimensions of an absorber and its proximity to the central regions of the host protogalaxy.

3.2 Column density distribution

For systems at $z > 1.6$, the Lyα absorption line is redshifted away from otherwise inaccessible UV wavelengths. For such systems, it is possible to determine observationally the column density of neutral hydrogen. We compiled a list of $N_{\rm H\,I}$-values for 38 GRB–DLAs in the redshift interval $2 < z < 4$ by combining data from Jakobsson et al. (2006), Chen et al. (2007) and Fynbo et al. (2009). We caution that, unlike for the QSO–DLAs for which homogeneous catalogues are available (e.g. Prochaska et al. 2005), our GRB–DLA statistics are formed from a heterogeneous sample. It is possible that extinction of UV/optical afterglows leads to some preference, in the observed data, to select lower column density absorbers (with correspondingly smaller dust columns; e.g. Fynbo et al. 2009); at this stage, however, the observational significance of this effect is not transparent – further discussion is given in Section 4.4.

Fig. 4 compares the observed column density distribution (points with error bars) with our simulated GRB–DLA column densities

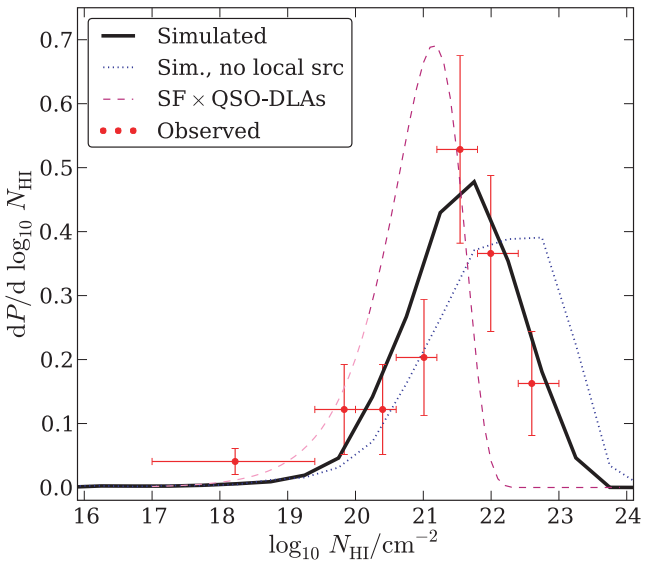

Figure 4. The differential column density distribution of GRBs as observed (points with error bars; see the text for references) and simulated (solid line); these are in agreement except for systems with $N_{\rm H\,I} < 10^{19}\,\mathrm{cm}^{-2}$. The dotted line shows the results obtained if stellar ionizing sources are ignored, leading to a slight overestimate of very high-column-density systems. The dashed line shows the results expected from a naive argument based on the observed QSO–DLA distribution combined with the Kennicutt star formation relation; this underestimates the incidence of high-column-density systems, which may be a useful result for understanding the high-redshift ISM (see the text for details).

(thick solid line). The form of the differential distribution is well matched by our simulations for $N_{\rm H\,I} > 10^{19}\,{\rm cm}^{-2}$. It is immediately clear that the GRB $N_{\rm H\,I}$ distribution is very different from the intervening case: for instance, the $N_{\rm H\,I}$ distribution diverges at low $N_{\rm H\,I}$ for intervening DLAs but approaches zero for GRB–DLAs; further the median GRB–DLA has a column density of $N_{\rm H\,I} \simeq 10^{21.5}\,{\rm cm}^{-2}$ whereas only 1 per cent of QSO–DLAs are so extreme in a catalogue derived from the Sloan Digital Sky Survey Data Release 5 (SDSS DR5) (Prochaska et al. 2005).[4] We will discuss these differences and their physical origins in Section 3.4.

Overall, the simulated column density distribution is in good agreement with the observed result. The median $N_{\rm H\,I}$ value ($N_{\rm H\,I} \simeq 10^{21.5}\,{\rm cm}^{-2}$) is reproduced quantitatively. The dotted line in Fig. 4 shows the output from our simulations without accounting for the effects of local sources (Section 2.1); in this case, a factor of 2 overabundance of very high column density sightlines ($N_{\rm H\,I} > 10^{22.5}\,{\rm cm}^{-2}$) is predicted. Our simple approach to accounting for the stellar UV emission has therefore mitigated this problem. However, our simulations produce a statistically significant shortfall of systems with $N_{\rm H\,I} < 10^{19}\,{\rm cm}^{-2}$ compared with the observed population; we will now discuss an estimate of the escape fraction from our simulations, which will highlight this difficulty.

3.3 Escape fraction

Since we have constructed a sample of neutral hydrogen column densities emerging from star formation regions, we can calculate the fraction of radiation at the Lyman limit which escapes relative to the total produced:

$$f_{\rm esc} = \int {\rm d}N_{\rm H\,I}\, p(N_{\rm H\,I})\, {\rm e}^{-\sigma_{\rm L} N_{\rm H\,I}}, \quad (4)$$

where $\sigma_{\rm L} = 6.28 \times 10^{-18}\,{\rm cm}^2$ is the photoionization cross-section of a hydrogen atom at the Lyman limit. In effect $f_{\rm esc}$ obeys almost identically $f_{\rm esc} = p(<1.6 \times 10^{17}\,{\rm cm}^{-2})$, since the exponential in the integral acts as a step function in log $N_{\rm H\,I}$ space; however, we evaluate the full integral since it is not a difficult computation. We ignore the effect of dust; although its rate of UV photon absorption can be comparable to that of the hydrogen in certain physical situations, Gnedin et al. (2008) argue plausibly that this has little impact on the escape fraction since most radiation leaks out along a small fraction of extremely low-column-density (essentially unobscured) sightlines.

Our calculated escape fraction, $f_{\rm esc} = 1.0$ per cent, is somewhat smaller than recent estimates have suggested. In particular, it falls short of the 2 ± 0.3 per cent derived by Chen et al. (2007) from reasonable extrapolation of the observed GRB–DLA statistics (see also Fynbo et al. 2009); this is a reflection of our previously described underestimate of the abundance of absorbers in the observational bin $10^{17} < N_{\rm H\,I}/{\rm cm}^{-2} < 10^{19}$. More worryingly, our derived fraction is very sensitive to assumptions we make about launching the GRB sightlines; for instance, adding a 200 pc Gaussian random offset to our stellar particle loci raises the fraction to $f_{\rm esc} = 2$ per cent. Such an approach might be justifiable if runaway massive stars are preferentially involved in GRB events (e.g. Hammer et al. 2006).

[4]The DR5 catalogue is available from www.ucolick.org/∼xavier/SDSSDLA/. Recently the DR7 catalogue was analysed with an independent pipeline (Noterdaeme et al. 2009); the differences are, however, minor for our present purposes.

While some adjustments can increase the reported escape fraction, others cause it to decrease. For instance, when we restricted sightlines to originate from younger stellar populations ($<10^7$ yr) than our fiducial technique imposes (Section 2.2), we derived an escape fraction of 0.7 per cent. We also implemented a model in which the GRB launch rate is made proportional to the instantaneous star formation rate as determined by the gas properties (i.e. taking $p_{\rm GRB} \propto \rho^{1.5}$ in regions where star formation is permitted, as opposed to probing new star particles formed by the simulation, Section 2.2). In this case, the escape fraction dropped to just 0.1 per cent. These changes appear to be because our stellar feedback algorithm rapidly decreases the density in regions around young stars in the live simulation, an effect underestimated in the fully instantaneous model – but perhaps overestimated in our standard approach which allows for excessive delays between formation and GRB events.

None of our other results – including the column density distribution for $N_{\rm H\,I} > 10^{19}\,{\rm cm}^{-2}$ – is strongly affected by these changes, so while the escape fraction seems to highlight shortcomings in our approximate schemes, we are confident that the majority of our results are robust. Further discussion of this point is given in Section 4.

3.4 Interpretation of the high $N_{\rm H\,I}$ distribution

As described above, both observed and simulated GRB–DLAs have an entirely different $N_{\rm H\,I}$-distribution from that of QSO–DLAs. To account for this, we should consider the relative importance of the two major distinctions between these populations (halo mass weighting and internal sightline positioning; Section 3.1). We find that the difference in the column density distributions arises chiefly from the GRB's propensity to probe inner regions of the halo rather than any difference in the column density distribution as a function of halo mass.

We verified this by (unphysically) weighting our QSO–DLA sightlines using the GRB–DLA halo weights ($w_{\rm h}$; equation 3) – the result was an increase in the median $N_{\rm H\,I}$ of just 0.1 dex, from $10^{20.5}$ to $10^{20.6}\,{\rm cm}^{-2}$. Conversely, by weighting our GRB–DLA sightlines using the QSO–DLA halo weights (based on the H I cross-section) we found that the median GRB–DLA column density decreased by just 0.1 dex. This result was foreshadowed by fig. 8 of P08 which shows that the mass of the hosting halo has relatively little effect on the $N_{\rm H\,I}$ value of a typical sightline.

The combination of an observationally derived fit to the QSO–DLA column density distribution (e.g. Prochaska et al. 2005) and the Schmidt–Kennicutt relation between gas surface density and star formation rate ($\dot{\Sigma}_\star \propto N_{\rm H\,I}^n$, $n \simeq 1.5$; Kennicutt 1998b) actually predicts a GRB–DLA column density distribution. This prediction is independent of the details of the underlying halo populations, relying only on the QSO absorbers to trace the unbiased neutral hydrogen distribution. If the column density distribution follows a Schechter function (Pei & Fall 1995; Prochaska et al. 2005) with asymptotic slope α and exponential scale N_γ, one has the following result for the distribution of GRB $N_{\rm H\,I}$ values:

$$\frac{{\rm d}P}{{\rm d}\log_{10} N_{\rm H\,I}} \propto N_{\rm H\,I}^{1+\alpha+n} \exp\left(\frac{-2N_{\rm H\,I}}{N_\gamma}\right), \quad (5)$$

where the 1 in the exponent arises from the change of variables $N_{\rm H\,I} \rightarrow \log_{10} N_{\rm H\,I}$ and the 2 in the exponential arises because, by symmetry, a sightline to the GRB intersects on average half of the total gas column density of the host galaxy.

We exhibit this function as a dashed line in Fig. 4, using the peak posterior dust-corrected[5] parameters (Pontzen & Pettini 2009) determined for the previously mentioned SDSS DR5 sample in the redshift interval $2.2 < z < 4$: $\alpha = -1.79, N_\gamma = 10^{21.61}\,\mathrm{cm}^{-2}$. Qualitatively, this approach predicts the correct form for the GRB–DLA $N_{\mathrm{H\,I}}$ distribution, with a peak and long tail to low column densities. However, quantitative details at high $N_{\mathrm{H\,I}}$ are in some tension with the observed GRB–DLA results. (One is less interested in pursuing exact agreement below $N_{\mathrm{H\,I}} < 10^{20.3}\,\mathrm{cm}^{-2}$ since such systems are not included in the fit to the quasar data.)

P08 showed that our simulated QSO–DLA column density distribution, while in good overall agreement with the observed distribution, overpredicts the occurrence of high $N_{\mathrm{H\,I}}$ quasar systems. Yet the present considerations show that the *observed* QSO–DLA distribution *under* predicts the occurrence of high $N_{\mathrm{H\,I}}$ GRB systems – while the simulations are in better agreement with observations! This three-way tension between the observed QSO–DLAs, observed GRB–DLAs and simulations may be a fruitful line of future exploration to understand details of the high-redshift ISM. For instance, high-$N_{\mathrm{H\,I}}$ quasar absorbers are plausibly limited by H_2 formation (Schaye 2001). The relative overabundance of high-$N_{\mathrm{H\,I}}$ absorbers in GRB afterglows might then suggest dissociation of molecular clouds in the vicinity of UV-bright young stellar populations responsible for the GRB events. Given the crudeness of our gas phase modelling, we are not in a position to make any certain conclusions at this stage.

We should take a moment to consider whether it is appropriate to apply the Kennicutt relation in the sense above. In a direct test of the relation's applicability to quasar DLA populations, Wolfe & Chen (2006) used the $N_{\mathrm{H\,I}}$ distribution to predict a background level of faint diffuse emission at rest-frame wavelengths $\lambda \simeq 1500$ Å. After blind searches in deep imaging, they concluded that the Kennicutt relation was overestimating star formation in DLAs. However, the authors also noted that this result may be negated if a substantial fraction of DLAs are associated with nearby bright compact sources, which were excluded from their search algorithms. Since this association indeed appears to hold in our simulations (P08), the relation may tentatively be applied. We note that the naive Kennicutt weighting applied to the simulated QSO–DLAs does produce the approximate distribution of simulated GRB–DLAs, so this method cannot constitute an entirely spurious comparison. Nagamine, Zhang & Hernquist (2008) also placed confidence in such a weighting, using it to adapt their simulated QSO–DLA statistics to discuss the expected redshift evolution of the GRB–DLA $N_{\mathrm{H\,I}}$ distribution.

3.5 Metallicity distribution

Each of our sightlines is assigned a neutral gas metallicity according to equation (2); using our fiducial techniques, we can then predict the metallicities observed in cosmological GRB–DLA samples. Fig. 5 shows the resulting distribution as a solid thick line, with the median indicated by a short vertical line bisecting the distribution. Our metallicity distribution peaks at around $\mathrm{Z}_\odot/3$, with a sharp decline to higher metallicities (and long tail to lower); as a result the median is approximately $\mathrm{Z}_\odot/6$.

As a rough observational comparison, we used the lists of Prochaska et al. (2007, 2009) and Ledoux et al. (2009) to compile eight GRB–DLA metallicity measurements and a further seven lower limits at $z > 2$; to our knowledge this is a complete census of such constraints. The small numbers, line saturation effects and unknown biases[6] make it hard, at present, to attempt a quantitative comparison of the observed and simulated metallicity distributions. We merely indicate in Fig. 5 the range of the observed metallicities (by a horizontal bar), with a vertical line indicating the median value. The shaded region shows the range of the median if we include assumed metallicity measurements of up to three times the value of the lower limits.

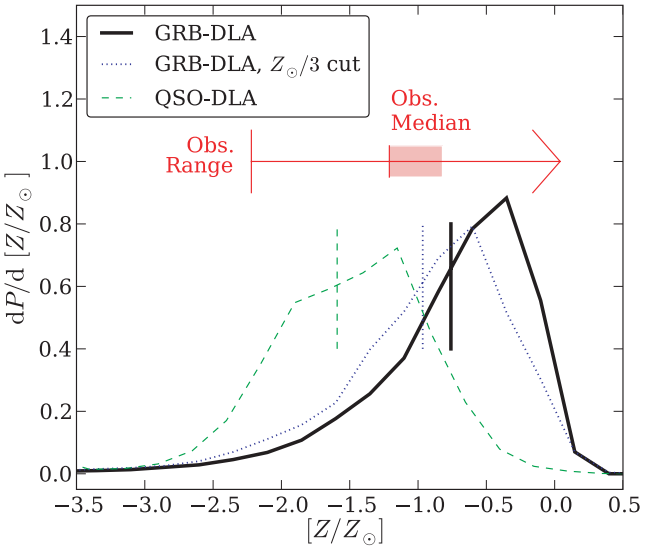

Figure 5. The neutral gas metallicity of our simulated GRB–DLA sightlines (solid line) compared with the same diagnostic of QSO–DLA sightlines (dashed line). GRB–DLAs have significantly higher metallicities as they are associated with star-forming regions of more massive haloes than QSO–DLAs. This picture is broadly supported by the handful of measurements of (and lower limits on) GRB neutral gas metallicities at $z > 2$, the range of which is shown by the horizontal bar. The shaded region shows the range of the median if we assume systems with line saturation have up to three times the metallicity of their measured lower limits (see the text for details). The dotted line shows the simulated results after imposing a restriction that GRBs can only arise in stellar progenitors with metallicity $Z < \mathrm{Z}_\odot/3$.

Given the uncertainties our simulations seem to provide an acceptable match. Although the observational constraints are not tight, there is no doubt that, for instance, our QSO–DLA metallicity distribution (shown as a dashed line with its median similarly indicated) would be deemed inconsistent with the data.[7]

Unlike for the column density distribution, we attribute much of this metallicity shift to the 1 dex higher halo masses probed by the typical GRB–DLA compared with a QSO–DLA (Section 3.1), in agreement with Fynbo et al. (2008). Mean neutral gas metallicity in our haloes range from $\sim 10^{-3}\,\mathrm{Z}_\odot$ for $M_{\mathrm{vir}} = 10^9\,\mathrm{M}_\odot$ to $\sim 10^{-1}\,\mathrm{Z}_\odot$ for $M_{\mathrm{vir}} = 10^{12}\,\mathrm{M}_\odot$: a 0.7 dex rise is expected from this effect. The remainder of the effect is accounted for by the marginally higher metallicities of the central regions which the GRB sightlines probe. This is a small $\sim$0.1 dex correction, despite large internal abundance

[5]The dust correction is small and does not in fact have a qualitative impact on the following considerations, but we include it to remove suspicion that dust biasing of QSO samples could account for any discrepancies.

[6]If anything, the observed sample is likely biased in favour of low-metallicity GRB–DLAs which have smaller dust columns and therefore brighter optical afterglows; see e.g. Fynbo et al. (2008).

[7]Our new simulations have a QSO–DLA metallicity distribution with a median $[Z/\mathrm{Z}_\odot] \simeq -1.6$, approximately 0.2 dex lower than the median in P08, as described in Section 2.

gradients discussed below, because the sightline averaging smears out much of the metal density variations.

3.6 Effect of GRB metallicity ceiling

So far we have assumed that GRB events simply trace, in an unbiased fashion, star formation activity in the cosmos. However, GRBs are thought to arise from rapidly rotating progenitor stars. Since angular momentum is concentrated in outer layers of high-mass stars, it is easily lost through winds. Mass-loss rates are thought to scale almost linearly with metallicity, so GRBs could well arise preferentially from low-metallicity progenitor stellar populations (Woosley & Bloom 2006, and references therein). There is also some observational evidence, particularly at low redshifts, that GRB-hosting populations are indeed biased in this sense (e.g. Fruchter et al. 2006; Modjaz et al. 2008; Calura et al. 2009; although see Wolf & Podsiadlowski 2007; Kelly, Kirshner & Pahre 2008; Savaglio, Glazebrook & LeBorgne 2009).

We tested the effect of such a picture on our results by producing a new set of sightlines, allowing GRBs to be launched only from stellar populations with $Z < \mathrm{Z}_\odot/3$ (as suggested by Woosley & Heger 2006) and recalculating the halo weights according to $w_\mathrm{h}' = w_\mathrm{h}\, f(Z < \mathrm{Z}_\odot/3)$, where w_h is defined by (3) and $f(Z < \mathrm{Z}_\odot/3)$ is the mass fraction of young stars (according to our criterion in Section 2.2) with metallicity below the adopted limit.

We found that the differences in column density and impact parameter distributions are entirely negligible; the only significant effect was seen in the sightline neutral gas metallicities and the stellar population metallicities (the latter quantity is discussed in Section 3.7). The updated neutral gas metallicity distribution is shown as a dotted line in Fig. 5 and shows a reduction in median metallicity by 0.2 dex to $[Z/\mathrm{Z}_\odot] = -0.9$. This is a relatively small shift, largely reflecting that even our most massive haloes at $z = 3$ have mean cold gas metallicities of only $\mathrm{Z}_\odot/2$.

Furthermore, sightlines of up to solar metallicity remain attainable with the cut because the neutral gas within each halo has metallicity dispersions of up to an order of magnitude; thus, sightlines which start in a low-metallicity environment may nonetheless encounter a nearby dense knot of solar metallicity gas which dominates the line-of-sight measurement. This is illustrated by Fig. 6, which shows the metallicity of the gas in a thin slice through the centre of the largest halo in the MG box. Red and white regions have metallicity such that they are excluded from hosting GRB progenitors. Overplotted contours show the regions in which the neutral gas volume density exceeds $n_{\mathrm{H\,I}} = 10^{-2}, 0.1, \ldots, 10^{2}\,\mathrm{cm}^{-3}$ (respectively from outer to innermost contour). Dense neutral regions ($>0.1\,\mathrm{cm}^{-3}$) have mean metallicity $10^{-0.4}\,\mathrm{Z}_\odot$ with a dispersion of 0.8 dex. Qualitatively GRBs are to be found in the densest regions which are not forbidden by the metallicity cut (i.e. not within the red area); the prime location remains very close to the centre so that high-metallicity gas will present a large cross-section to the sightline.

3.7 GRB host halo properties

Our simulation naturally associates each QSO or GRB sightline with a host halo, the properties of which can readily be computed. In this section, we briefly describe the distribution of star formation rates, total stellar masses and stellar metallicities in the populations within each halo.

Fig. 7 shows the distribution of each of these properties (in three separate panels) with the GRB–DLAs and QSO–DLAs represented by solid and dashed lines, respectively. The dotted lines show

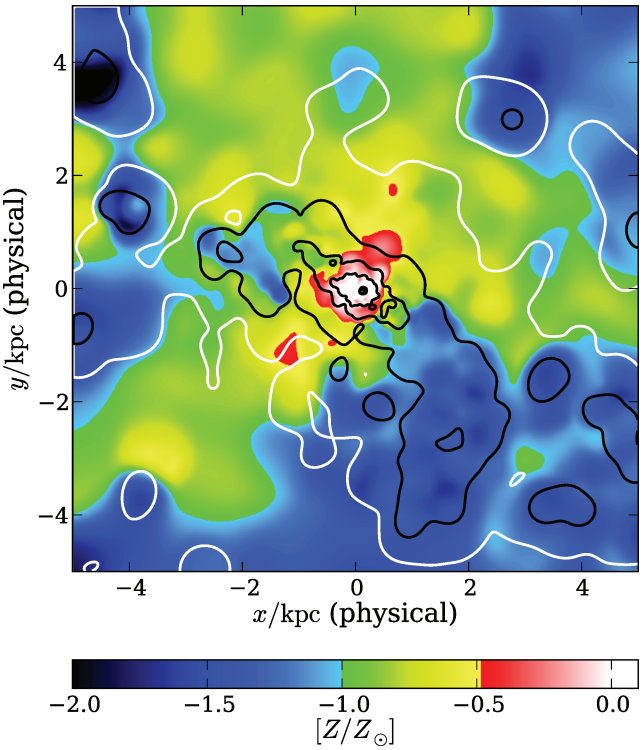

Figure 6. Map of the metallicity of the gas in a thin slice through the centre of the most massive halo in our MG box ($M_\mathrm{vir} \simeq 3 \times 10^{10}\,\mathrm{M}_\odot$), with the volume density of H I overplotted as contours (the outermost white contour corresponds to $n_{\mathrm{H\,I}} = 10^{-2}\,\mathrm{cm}^{-3}$ and inner, black, contours indicate 0.1, 1, 10 and $10^2\,\mathrm{cm}^{-3}$). Little of the gas by volume exceeds 1/3 solar (red region, the cut-off in our truncated GRB model), and star formation in nearby regions can still give rise to GRBs with lines of sight intersecting the high-Z region.

the population of GRB–DLAs if a progenitor metallicity criterion $Z < \mathrm{Z}_\odot/3$ is introduced (see Section 3.6). The GRB–DLA hosts generally show somewhat higher star formation rates, total stellar masses and metallicities than the intervening population: mean star formation rates[8] for the GRB and QSO DLAs are 10 and 1 $\mathrm{M}_\odot\,\mathrm{yr}^{-1}$, a difference of 1 dex. However, this should be compared with a wide dispersion of around 4 dex in both cases. We have added to the top panel a scale showing corresponding unobscured UV magnitudes using the luminosity conversion ratio from Kennicutt (1998a), corrected by a factor of 1.6 in order to consistently use the Kroupa IMF assumed by our simulations.

Mean stellar masses for the QSO–DLA and GRB–DLA host populations differ by around an order of magnitude (with respective values $10^{8.9}$ and $10^{9.9}\,\mathrm{M}_\odot$), but exhibit wide dispersions around these values. For the stellar metallicity (distinct from the neutral gas metallicities described in Section 3.5), the mean values are $10^{-0.7}$ and $10^{-0.9}\,\mathrm{Z}_\odot$ for unbiased and metallicity-biased GRBs, respectively, compared against $10^{-1.4}\,\mathrm{Z}_\odot$ for the QSO DLAs.

The agreement of these results with observational constraints has been discussed elsewhere for QSOs (P08). For GRBs, no homogeneous high-redshift catalogues of host galaxies have been assembled. However, the four $z \gtrsim 1$ objects studied by Christensen, Hjorth

[8]Calculated as described in P08 section 4.3. The QSO distribution shown in the middle panel of Fig. 7 is equivalent to fig. 14 of P08; however, the latter plot shows $\mathrm{d}p/\mathrm{d}\dot{M}_\star$, not $\mathrm{d}p/\mathrm{d}\log_{10}\dot{M}_\star$ as its label in the published version erroneously claims – this was a typographical error.

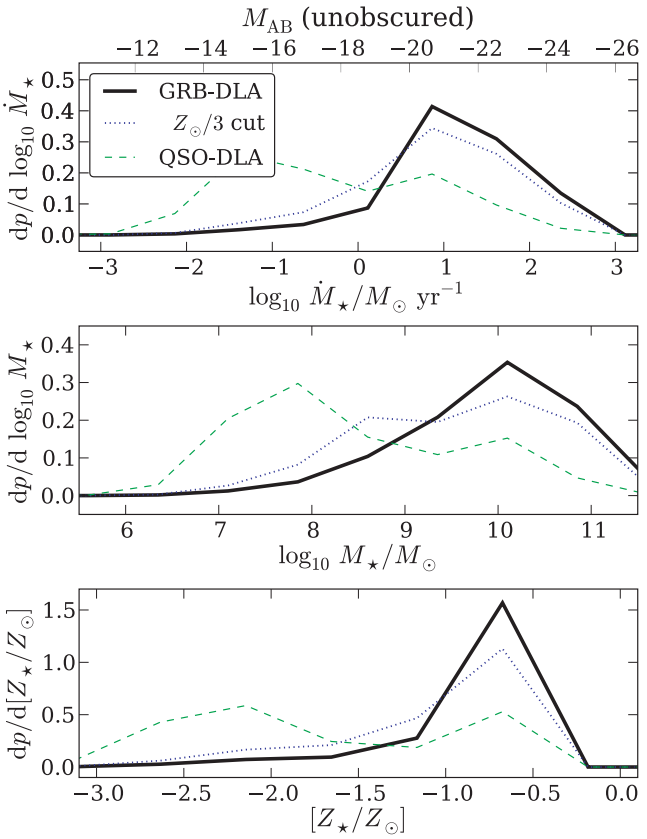

Figure 7. Properties of the stellar population within the host galaxies of $z = 3$ GRB–DLAs (thick solid line, dotted line for $<Z_\odot/3$ progenitors) and QSO–DLAs (dashed line). From top to bottom, we plot the probability density distributions for star formation rate, total stellar mass and (mass-weighted) stellar population metallicity, respectively. A second scale on the uppermost panel converts the star formation rate to an UV AB magnitude as described in the text, ignoring dust obscuration. The properties of the host galaxies of the two classes of absorber differ in accordance with their different virial masses, with the GRBs favouring somewhat higher star formation rates, total stellar masses and stellar metallicities. This applies even for the cut metallicity GRB population, although the high-Z peak is slightly suppressed in this case.

& Gorosabel (2004) have star formation rates well within the range suggested by our simulations; see also Jakobsson et al. (2005). At $z > 2$, the eight GRB–DLAs with identified host galaxies listed by Fruchter et al. (2006) have a range of luminosities $-16 \geq M_V \geq -22$ in the observer-frame V band; broadly, given that our brightest objects should be substantially dust obscured in the rest-frame UV, our results are consistent with these observations.

All of the distributions can be interpreted by inspecting the relation between virial mass and the property under consideration (see e.g. fig. 13 of P08 for the star formation rate). As well as satisfactorily accounting for the approximate means in each case, specific features in the distribution function can be explained in this way. In particular, $Z(M_{\rm vir})$ rises steeply from $10^{-3}\ \mathrm{Z}_\odot$ at $M_{\rm vir} = 10^9\ \mathrm{M}_\odot$ to around half solar at $10^{11}\ \mathrm{M}_\odot$, beyond which the metallicity remains constant. Physically, this is because massive haloes harbour older stellar populations as well as the young enriched stars from recent formation activity. This gives rise to bimodality in the metallicity of the QSO–DLA stellar populations, with peaks at $Z \sim 10^{-2.0}\ \mathrm{Z}_\odot$ (corresponding to the metallicity of the peak QSO–DLA mass, $10^{10}\ \mathrm{M}_\odot$) and $Z \sim 10^{-0.5}\ \mathrm{Z}_\odot$ (corresponding to the plateau metallicity: all haloes with $M > 10^{11}\ \mathrm{M}_\odot$ fall into this bin). Bimodality is not seen in the neutral gas metallicities because these measurements are averaged only along the line of sight, not the entire halo, and local metallicities can reach values of up to solar (see Fig. 6).

4 CONCLUSIONS AND DISCUSSION

4.1 Overview

We have investigated the properties of high-redshift ($z \sim 3$) GRB–DLA absorbers in a series of galaxy formation simulations, on the assumption that long-duration GRBs are associated with the death of massive, short-lived stars. Following our success in matching properties of traditional DLAs along randomly chosen sightlines to quasars (P08), we used our results to interpret observational differences between the populations, finding two primary distinctions.

First, because the star formation rate scales more steeply with virial mass ($\dot{M}_\star \sim M_{\rm vir}^{1.6}$; P08, fig. 13) than does the neutral hydrogen cross-section (approximately $\sigma_{\rm DLA} \sim M_{\rm vir}$, although a single power law is an oversimplification; P08, fig. 4) the haloes contributing most to the GRB–DLA incidence rate are considerably more massive than those contributing to the QSO–DLA incidence rate (Section 3.1). The dominant contribution for GRB–DLAs arises in haloes with $10^{10} < M_{\rm vir}/\mathrm{M}_\odot < 10^{12}$; for QSO–DLAs, this range is an order of magnitude lower, $10^9 < M_{\rm vir}/\mathrm{M}_\odot < 10^{11}$. While this argument obviously relies on our simulated scalings being accurate, there are many reasons for being cautiously optimistic: in particular, the simultaneous matching of the QSO–DLA metallicity distribution (P08) and the galaxy mass–metallicity relation (Brooks et al. 2007) is hard to reproduce unless the underlying relations are realistic. Fynbo et al. (2008) have previously reported that GRB metallicities can be accounted for if a Holmberg-like relation, $\sigma_{\rm DLA} \propto \dot{M}_\star^{2t}$ with $t \simeq 0.4$, holds at high redshift; here, we have shown that the metallicities are reproduced and that just such a scaling arises naturally from cosmological simulations (albeit with $t \sim 0.3$). The difference in mass ranges has some effect on the predicted properties of the hosting haloes, most notably lifting the mean stellar metallicity from $10^{-1.4}\ \mathrm{Z}_\odot$ for QSO–DLAs to $10^{-0.7}\ \mathrm{Z}_\odot$ for GRB–DLAs. However, these effects are relatively small compared to the large spread of hosting halo properties in both populations, reflecting that incidence rates for both classes are a shallow function of virial mass.

Courty, Björnsson & Gudmundsson (2004) took a reciprocal approach, using observed galaxy properties to determine the likely underlying population of host objects. Their results showed that GRB hosts at $z = 3$ have typical masses $M_{\rm vir} < 10^{10}\ \mathrm{M}_\odot$ which does not agree with our analysis. This may be because Courty et al. (2004) adopt, on an observational basis, the star formation efficiency (rate of star formation per unit stellar mass) as an indicator of GRB probability. Our simulations do not suggest a strong link between star formation efficiency defined in this way and the GRB event rate; if extended observational analysis continues to point to such a link, further investigation will be required.

The second distinguishing effect we find is that the GRB–DLAs arise in gas closer to the centres of the hosting objects than a conventional DLA (as seen observationally; e.g. Savaglio, Fall & Fiore 2003; Vreeswijk et al. 2004; Fruchter et al. 2006). While resolution artefacts cloud an exact interpretation of some details, we are confident in concluding that approximately 50 per cent of GRB–DLAs arise less than 1 kpc from the centre of their parent halo, whereas this figure for QSO–DLAs is just 10 per cent (and, accordingly, the

QSO–DLAs have a higher median impact parameter of ∼4 kpc). Similarly, GRB–DLAs tend to be more concentrated along the line of sight; we defined the measure Δ to be the distance over which half the total hydrogen column density is accumulated and found that this took a median value of 1 kpc for GRB–DLAs and 2 kpc for QSO–DLAs. This indicates that the neutral hydrogen seen in GRB afterglow spectra is not to be found in the immediate circumstellar environment, but is spread over an extended region – although a detailed analysis would require better resolution of the ISM than our present simulations are able to provide; see Section 4.4.

4.2 Observational diagnostics

We investigated the neutral hydrogen column density (Sections 3.2, 3.4) and the metallicity (Section 3.5) distributions of our simulated absorbers, finding good observational agreement in both cases except for a deficiency in low-column-density ($N_{\rm H\,I} < 10^{19}\,{\rm cm}^{-2}$) absorbers. The simulated neutral hydrogen column densities, as for the observed GRB absorbers, are much higher than those seen in intervening DLAs. We conclude that this is largely due to the lowered impact parameters, with the change in the masses of haloes having a secondary effect (for instance, fig. 8 of P08 shows that the relation between column density and halo mass is relatively weak).

The metallicities are higher, by almost a factor of 10, than those seen in QSO–DLAs – even if GRB–DLAs can only arise from stellar populations with $Z < {\rm Z}_{\odot}/3$ (a limit motivated by the collapsar picture, e.g. Hirschi, Meynet & Maeder 2005; Woosley & Heger 2006). This suggests that the observed population of high-redshift GRB–DLAs and GRB hosts do not pose any severe problems for the typical collapsar GRB paradigms, counter to the conclusions of Wolf & Podsiadlowski (2007) and Lapi et al. (2008), but in agreement with Nuza et al. (2007). The ineffectiveness of a metallicity ceiling on the GRB–DLA progenitors to limit the metallicity of the overall sightline population arises because, within each protogalaxy, neutral gas metallicity dispersions of up to an order of magnitude are possible (greatly exceeding the 0.2 dex allowed by Wolf & Podsiadlowski 2007). In our picture, GRBs now originate in dense regions with lower metallicities, but they are typically still close to gas with significantly higher metallicity (Fig. 6) which consequently presents a large covering fraction. Furthermore, high-mass, metal-rich haloes continue to contribute significantly to the GRB rate: even for our most massive galaxy at $z = 3$ (with a mean neutral gas metallicty of $\simeq {\rm Z}_{\odot}/2$) 30 per cent of the star-forming gas mass remains at low metallicities $Z < {\rm Z}_{\odot}/3$. We note that, since our simulations include a turbulent metal diffusion term, this dispersion cannot be attributed solely to numerical effects (Section 2).

The leaves open the question of whether the observational situation at lower redshifts ($z \lesssim 1.5$), where host galaxies may be less massive than typical hosts of supernovae (Fruchter et al. 2006; although see Kelly et al. 2008; Savaglio et al. 2009), can be quantitatively reconciled with less pronounced differences between supernovae and GRB hosts at high redshift. A natural explanation is simply that at earlier times mean metallicities are lower and high-metallicity gas is typically less well mixed (Nuza et al. 2007; Kocevski, West & Modjaz 2009). In principle, we could probe this claim in our simulations, but such work is beyond the scope of the present investigation.

4.3 UV escape fraction

One of the shortcomings of our work is the unreliability of our predicted 1 per cent escape fraction of UV photons from GRB regions. This is somewhat lower than recent simulations of escape fractions (∼2 per cent from results by Gnedin et al. 2008; between 2 and 10 per cent according to Razoumov & Sommer-Larsen 2006; or, for haloes with masses $M \lesssim 10^9\,{\rm M}_{\odot}$, Wise & Cen 2009 reported escape fractions between 25 and 80 per cent in very high resolution Eulerian simulations). Furthermore, it falls short of constraints on escape fractions from comparing rest-frame UV emission on either side of the Lyman limit break (∼2 per cent, although this value is sensitive to assumptions about the dust extinction; Shapley et al. 2006) and most directly underpredicts the escape fraction implied by current GRB observations (∼2 per cent; Chen et al. 2007, also visible directly in Fig. 4). However, most importantly, it is not robust to changes in our sightline construction algorithms (Section 3.2), unlike our other reported results.

The fragility reflects enormous uncertainties in this regime. In addition to the inherent complexities in accurately simulating absorption systems, for GRB–DLAs one can no longer justify the assumption that local ionizing UV sources are unimportant compared with cosmological sources (Section 2.1). We have taken a simple pragmatic approach to this problem by noting that the ionizing radiation of local sources is largely absorbed by H II regions which we paint into our output. (More accurate radiative transfer would probably lend confidence to our results, but we caution that without properly resolved H II regions the physical meaning of even the most precise transfer algorithms is still obscure.) Furthermore, there is a sensitivity to the energy injection from supernovae, as demonstrated by the drop in escape fraction to 0.1 per cent when triggering GRBs based on the star formation rate scaling of our gas particles (rather than existing young stellar particles which will have injected feedback energy into the nearby ISM). Finally, if runaway stars are important contributors to the GRB incidence rate (Hammer et al. 2006) one could expect a significant increase in the number of low column density absorbers.

4.4 Model uncertainties

There are further complications specific to GRB–DLAs which we have not specifically addressed. One is the evidence, from fine-structure transitions, that the neutral-phase metals detected in the systems are found at distances ranging from hundreds of parsecs (Prochaska et al. 2006) to as much as a kiloparsec or more (Vreeswijk et al. 2007; D'Elia et al. 2009). Given the known afterglow curves, the ionization of material by the GRB itself has been advanced as a possible explanation (e.g. Vreeswijk et al. 2007), although whether bubbles of more than a few parsecs radius can be ionized is unclear (e.g. Prochaska et al. 2008). The afterglow ionizing scenario is certainly not firmly established: perhaps the hot progenitor stellar populations have already carved out large ionized bubbles within the ISM, or the densest knots of gas which are metal rich are simply found at some distance from the GRB site (as we have mentioned above). Unfortunately we do not have resolution to probe the required scales reliably, but we did produce comparison statistics in which we ionized all material within a 1 kpc sphere of a GRB site before measuring its sightline properties. This led to a decrease in the numbers of very high column density absorbers ($N_{\rm H\,I} > 10^{22}\,{\rm cm}^{-2}$), reducing the median to $N_{\rm H\,I} = 10^{20.9}\,{\rm cm}^{-2}$. While the agreement of our simulations with observations was therefore impaired, our results were not qualitatively altered. Simulations with better resolution of the ISM will be necessary to properly constrain the effects of GRBs on their immediate environment.

Similarly, we have been unable to make a quantitative assessment of the possible biasing effects of dust. It is possible that dust obscuration of optical afterglows significantly biases against observing high metallicity, high column density absorbers. Indeed, the phenomena of 'dark bursts', GRBs for which the optical afterglow is not seen or is dimmed by several magnitudes, could be explained if these bursts are simply obscured by large columns of dust (Perley et al. 2009). However, it appears to be inappropriate to extrapolate the dust columns predicted for quasar absorbers (by models such as those in Pontzen & Pettini 2009) to our simulated GRB results: such an approach would significantly overpredict the reddening seen in real GRB–DLA samples (for a compilation, see Fynbo et al. 2009) and would suggest, for example, that GRB 080607 (with $N_{\rm H\,I} = 10^{22.7}\,{\rm cm}^{-2}$, $Z \sim {\rm Z}_{\odot}$ according to Prochaska et al. 2009) should have been completely dark. Either the GRB–DLAs arise in regions in which simple scalings of dust columns fail or perhaps the intense outgoing radiation is able to rapidly destroy dust even beyond the hydrogen ionization front. Whatever the explanation, we have been unable to make full sense of this issue in our present exploration and leave its resolution to a future work.

Despite these difficulties, we have affirmed in the present work that GRB–DLAs offer a new insight into the inner regions of the protogalactic population at high redshifts. Their basic statistics can be matched by simple techniques applied to existing cosmological simulations. There can be little doubt that, as observational data sets increase in size and sophistication, GRBs and their associated absorbers will make a significant contribution to our understanding of the structure and formation of galaxies.

ACKNOWLEDGMENTS

We acknowledge helpful conversations with Jason X. Prochaska, Susanna Vergani and Lise Christensen. AP is supported by an STFC studentship and scholarship at St John's College, Cambridge. FG acknowledges support from grants *HST* GO-1125, NSF ITR grant PHY-0205413, NSF AST-0607819 and NASA ATP NNX08AG84G. We thank the computer resources and technical support by TERAGRID, ARSC, NAS and the UW computing center, where the simulations were run.

REFERENCES

Brooks A. M., Governato F., Booth C. M., Willman B., Gardner J. P., Wadsley J., Stinson G., Quinn T., 2007, ApJ, 655, L17
Calura F., Dessauges-Zavadski M., Prochaska J. X., Matteucci F., 2009, ApJ, 693, 1236
Chen H.-W., Prochaska J. X., Gnedin N. Y., 2007, ApJ, 667, L125
Christensen L., Hjorth J., Gorosabel J., 2004, A&A, 425, 913
Courty S., Björnsson G., Gudmundsson E. H., 2004, MNRAS, 354, 581
Crain R. A. et al., 2009, MNRAS, 399, 1773
D'Elia V. et al., 2009, ApJ, 694, 332
Dunkley J. et al., 2009, ApJS, 180, 306
Fruchter et al., 2006, Nat, 441, 463
Fynbo J. P. U. et al., 2006, A&A, 451, L47
Fynbo J. P. U., Prochaska J. X., Sommer-Larsen J., Dessauges-Zavadsky M., Møller P., 2008, ApJ, 683, 321
Fynbo J. P. U. et al., 2009, ApJS, 185, 526
Gehrels N. et al., 2004, ApJ, 611, 1005
Gnedin N. Y., Kravtsov A. V., Chen H.-W., 2008, ApJ, 672, 765
Governato F., Willman B., Mayer L., Brooks A., Stinson G., Valenzuela O., Wadsley J., Quinn T., 2007, MNRAS, 374, 1479
Governato F. et al., 2009, MNRAS, 398, 312
Haardt F., Madau P., 1996, ApJ, 461, 20
Hammer F., Flores H., Schaerer D., Dessauges-Zavadsky M., Le Floc'h E., Puech M., 2006, A&A, 454, 103
Hirschi R., Meynet G., Maeder A., 2005, A&A, 443, 581
Jakobsson P. et al., 2005, MNRAS, 362, 245
Jakobsson P. et al., 2006, A&A, 460, L13
Katz N., White S. D. M., 1993, ApJ, 412, 455
Kelly P. L., Kirshner R. P., Pahre M., 2008, ApJ, 687, 1201
Kennicutt R. C., Jr, 1998a, ARA&A, 36, 189
Kennicutt R. C., Jr, 1998b, ApJ, 498, 541
Kocevski D., West A. A., Modjaz M., 2009, ApJ, 702, 377
Kroupa P., Tout C. A., Gilmore G., 1993, MNRAS, 262, 545
Lapi A., Kawakatu N., Bosnjak Z., Celotti A., Bressan A., Granato G. L., Danese L., 2008, MNRAS, 386, 608
Ledoux C., Petitjean P., Srianand R., 2003, MNRAS, 346, 209
Ledoux C., Vreeswijk P. M., Smette A., Fox A. J., Petitjean P., Ellison S. L., Fynbo J. P. U., Savaglio S., 2009, A&A, 506, 661
Lodders K., 2003, ApJ, 591, 1220
MacFadyen A. I., Woosley S. E., 1999, ApJ, 524, 262
Maiolino R. et al., 2008, A&A, 488, 463
Mirabal N. et al., 2002, ApJ, 578, 818
Modjaz M. et al., 2008, AJ, 135, 1136
Nagamine K., Zhang B., Hernquist L., 2008, ApJ, 686, L57
Noterdaeme P., Petitjean P., Ledoux C., Srianand R., 2009, A&A, 505, 1087
Nuza S. E., Tissera P. B., Pellizza L. J., Lambas D. G., Scannapieco C., de Rossi M. E., 2007, MNRAS, 375, 665
Osterbrock D. E., Ferland G. J., 2006, Astrophysics of Gaseous Nebulae and Active Galactic Nuclei, 2nd edn. University Science Books, Sausalito, CA
Pei Y. C., Fall S. M., 1995, ApJ, 454, 69
Perley D. A. et al., 2009, AJ, 138, 1690
Perna R., Loeb A., 1998, ApJ, 501, 467
Pettini M., 2006, in LeBrun V., Mazure A., Arnouts S., Burgarella D., eds, The Fabulous Destiny of Galaxies: Bridging Past and Present. Frontier Group, Paris, p. 319
Pontzen A., Pettini M., 2009, MNRAS, 393, 557
Pontzen A. et al., 2008, MNRAS, 390, 1349
Prochaska J. X., Herbert-Fort S., Wolfe A. M., 2005, ApJ, 635, 123
Prochaska J. X., Chen H.-W., Bloom J. S., 2006, ApJ, 648, 95
Prochaska J. X., Chen H.-W., Dessauges-Zavadsky M., Bloom J. S., 2007, ApJ, 666, 267
Prochaska J. X., Dessauges-Zavadsky M., Ramirez-Ruiz E., Chen H.-W., 2008, ApJ, 685, 344
Prochaska J. X. et al., 2009, ApJ, 691, L27
Razoumov A. O., Sommer-Larsen J., 2006, ApJ, 651, L89
Read J. I., Pontzen A. P., Viel M., 2006, MNRAS, 371, 885 (P08)
Savaglio S., 2006, New Journal of Physics, 8, 195
Savaglio S., Fall S. M., Fiore F., 2003, ApJ, 585, 638
Savaglio S., Glazebrook K., LeBorgne D., 2009, ApJ, 691, 182
Schaye J., 2001, ApJ, 562, L95
Shapley A. E., Steidel C. C., Pettini M., Adelberger K. L., Erb D. K., 2006, ApJ, 651, 688
Shen S., Wadsley J., Stinson G., 2009, MNRAS, submitted (arXiv:0910.5956)
Stinson G., Seth A., Katz N., Wadsley J., Governato F., Quinn T., 2006, MNRAS, 373, 1074
Tescari E., Viel M., Tornatore L., Borgani S., 2009, MNRAS, 397, 411
Thielemann F.-K., Nomoto K., Yokoi K., 1986, A&A, 158, 17
Tumlinson J., Prochaska J. X., Chen H.-W., Dessauges-Zavadsky M., Bloom J. S., 2007, ApJ, 668, 667
Vázquez G. A., Leitherer C., 2005, ApJ, 621, 695
Vreeswijk P. M. et al., 2004, A&A, 419, 927
Vreeswijk P. M. et al., 2007, A&A, 468, 83
Wadsley J. W., Stadel J., Quinn T., 2004, New Astron., 9, 137
Weaver T. A., Woosley S. E., 1993, Phys. Rep., 227, 65

Wise J. H., Cen R., 2009, ApJ, 693, 984
Wolf C., Podsiadlowski P., 2007, MNRAS, 375, 1049
Wolfe A. M., Chen H.-W., 2006, ApJ, 652, 981
Wolfe A. M., Gawiser E., Prochaska J. X., 2005, ARA&A, 43, 861
Woosley S. E., Bloom J. S., 2006, ARA&A, 44, 507
Woosley S. E., Heger A., 2006, ApJ, 637, 914

This paper has been typeset from a TEX/LATEX file prepared by the author.